Library and Information Science Research 39 (2017) 224–233

Contents lists available at ScienceDirect

Library and Information Science Research

journal homepage: www.elsevier.com/locate/lisres

Social scientists' data reuse behaviors: Exploring the roles of attitudinal beliefs, attitudes, norms, and data repositories

Ayoung Yoon[a], Youngseek Kim[b,*]

[a] Indiana University-Purdue University, School of Informatics and Computing, Department of Library and Information Science, 535 W. Michigan St. IT563, Indianapolis, IN 46202, United States

[b] University of Kentucky, School of Information Science, 331 Little Library Building, Lexington, KY 40506, United States

ABSTRACT

Many disciplines within the social sciences have a dynamic culture of sharing and reusing data. Because social science data differ from data in the hard sciences, it is necessary to explicitly examine social science data reuse. This study explores the data reuse behaviors of social scientists in order to better understand both the factors that influence those social scientists' intentions to reuse data and the extent to which those factors influence actual data reuse. Using an integrated theoretical model developed from the theory of planned behavior (TPB) and the technology acceptance model (TAM), this study provides a broad explanation of the relationships among factors influencing social scientists' data reuse. A total of 292 survey responses were analyzed using structural equation modeling. Findings suggest that social scientists' data reuse intentions are directly influenced by the subjective norm of data reuse, attitudes toward data reuse, and perceived effort involved in data reuse. Attitude toward data reuse mediated social scientists' intentions to reuse data, leading to the indirect influence of the perceived usefulness and perceived concern of data reuse, as well as the indirect influence of the subjective norm of data reuse. Finally, the availability of a data repository indirectly influenced social scientists' intentions to reuse data by reducing the perceived effort involved.

1. Introduction

There is a long tradition of sharing and reusing data in the social sciences. Hedrick (1988) argues that data sharing has been a concern for researchers since the late 1970s. However, while there were (and are) difference within disciplines, discussions about the value and sharing of social science data began in the early 1960s (Clubb, Austin, Gedda, & Traugott, 1985). For decades, the topic has intrigued researchers working with large-scale survey data, archivists at institutional repositories, and individuals who were frustrated with unsuccessful attempts to obtain other researchers' data. Fear (2013) asserts that this tradition of sharing and reusing data in the social sciences is due to the nature of social research, which often requires large amounts of unique data collected over time.

While there is no agreed upon formal definition of "social science data", the term has been generally understood to mean "numeric files originating from social research methodologies or administrative records, from which statistics are produced" (Inter-university Consortium for Political and Social Research [ICPSR], 2016). As implied by this definition, quantitative data have been the dominant form of data in social science, and Fear (2013) states that reuse of such data from repositories is the most common type of data reuse in social science. Other types of data have been also generated and reused in social science; for instance, qualitative data reuse is an established practice in some social science disciplines (Yoon, 2014b) and discussions of qualitative data sharing and reuse have emerged in journals such as *Forum: Qualitative Social Research* (Bergman & Eberle, 2005) and *IASSIST Quarterly* (Rasmussen, 2010).

While the social sciences, broadly speaking, have had a dynamic culture of sharing and reusing data, much of the research on data reuse in recent years has focused primarily on the life and physical sciences. Social science data differ from the data from lab-based or other life and physical science research. Social science data typically involve observations about human subjects and unstructured formats (e.g., interview transcripts, observation notes, and survey data). The data practices of social science research are arguably different as well; because they involve human subjects, they are usually regulated by institutional review boards. Understanding and interpreting the unstructured data collected in the social sciences often requires detailed contextual information. Given the breadth and importance of the

* Corresponding author.
E-mail addresses: ayyoon@iupui.edu (A. Yoon), youngseek.kim@uky.edu (Y. Kim).

http://dx.doi.org/10.1016/j.lisr.2017.07.008
Received 1 September 2016; Received in revised form 28 March 2017; Accepted 18 July 2017

differences of social science data, its reuse cannot be understood without studying it explicitly and exclusively from life and physical science data.

2. Problem statement

Several recent studies have investigated data reuse practices and behaviors in the social sciences (Daniels, Faniel, Fear, & Yakel, 2012; Faniel, Kriesberg, & Yakel, 2012; Faniel, Kriesberg, & Yakel, 2016; Niu, 2009; Yoon, 2014b, 2016, 2017) as part of a broader drive to understand data practices within the social sciences. While these studies have captured some of the contexts and characteristics of social science data reuse by focusing on specific aspects of data reuse practices, which may overlap with other disciplinary contexts, fewer studies have used theoretical approaches or models to explain social scientists' data reuse behaviors.

A theoretical model is often used to explain the meaning, nature, and challenges associated with the phenomena of interest, and it helps scholars to understand these phenomena more effectively. The lack of a theoretical model in the data reuse research leaves a significant gap in our understanding of the way disciplinary, organizational, and individual characteristics interact to encourage or discourage data reuse. This study advances a theoretical model of social scientists' data reuse behaviors. Specifically, this study explores the factors that influence social scientists' intentions to reuse data and the extent to which those factors influence actual data reuse. The theoretical model provides a broad explanation of the relationships that exist between factors that influence social scientists' data reuse. The conceptual underpinnings of this study will provide a new perspective for understanding data reuse behaviors, and will contribute to both theory and practice.

Although data reuse is important for many natural science and engineering disciplines, data reuse is becoming increasingly significant in social sciences in the context of data-intensive research as researchers reuse shared data sets along newly collected data sets. A better understanding of data reuse in the social sciences can help social scientists to compare their studies to existing ones and conduct more advanced studies based on shared and accumulated data sets. This study can also offer valuable insights for academic libraries seeking to develop or improve data stewardship services by taking into account the diverse factors affecting social scientists' data reuse behaviors.

3. Literature review

Many researchers discuss the social and individual benefits of data reuse. Data reuse expands research possibilities and saves on data (re) collection costs (Borgman, 2012). Yoon (2015), based on empirical research targeting social scientists, reported the perceived benefits of and motivations for using existing data, which included the data reusers' awareness of the usefulness of secondary data, the cost-effectiveness of reusing data, the ability to use large sample data, and the expediency of reusing data for training and education. Although there are some common benefits and motivations reported by others for all data reuse, Yoon (2015) found that the reuse of data gathered from large samples can be particularly helpful in verifying and generalizing prior findings in quantitative social science research. Curty (2015) also found that social scientists' data reuse intentions are mainly affected by perceived benefit involved in data reuse.

Despite the potential benefits of data reuse, many social scientists still have concerns, and it is known that they have more concerns about qualitative data reuse than about other types of data reuse. Bishop (2009) reported that qualitative researchers expressed concerns about potential ethical violations, since qualitative research involves direct interaction with human subjects. In addition, although the possibility of misinterpretation is a concern with all data reuse, qualitative researchers are more concerned with the nature of their data in general, because "knowledge about qualitative data is highly contextual and experience-dependent" (Niu & Hedstrom, 2008, p. 7). Reused data can also be perceived as less valuable (Goodwin, 2012; Martin, 1995), and the qualitative researchers in Yoon's (2014b) study faced challenges publishing their work which reused existing data; this too raised concerns about reusing data. Curty (2015) also reported that perceived risk involved in data reuse significantly affected social scientists' data reuse intentions.

Discovering relevant data may be challenging for scientists across disciplines (e.g., Faniel & Jacobsen, 2010; Zimmerman, 2008), but it is especially difficult for social scientists because data are distributed among various sources and systems (Yoon, 2015). Easy access to data was one of the most influential factors in determining social scientists' satisfaction with data reuse (Faniel, Kriesberg, & Yakel, 2016). Data repositories have a long history in social science, and they are known to support easy access to and reuse of available data through value-adding activities (Daniels, Faniel, Fear, & Yakel, 2012; Yoon, 2014a). However, social scientists searched for more data than was deposited in the repositories. In addition, Curty (2015) found that social scientists' data reuse is influenced by facilitating conditions such as documentation, repository, support, and training.

Even when data reusers can find sufficient, seemingly suitable data, data reuse can still pose challenges. Data reusers need to assess data before reusing it because they are usually unfamiliar with the details of the data. Reusers assess data for a good fit for the purpose of their study (Faniel, Kansa, Kansa, Barrera-Gomez, & Yakel, 2013), for data quality (Cragin & Shankar, 2006; Van House, 2002), or generally for reusability (Faniel & Jacobsen, 2010). Social scientists are also concerned with choosing good quality, trustworthy data and avoiding data with errors (Yoon, 2014a, 2016, 2017). Assessing data for each of these qualities requires different criteria; some important assessment factors which have been identified include data producers' ability to generate trustworthy data, other reusers' positive experiences using the data, and soundness of methodology used to produce data (Faniel & Jacobsen, 2010; Faniel, Kansa, Kansa, Barrera-Gomez, & Yakel, 2013; Yoon, 2017; Zimmerman, 2008).

A particular challenge arises from the fact that reusers have not participated in the initial study design and data collection process; thus, it can be difficult for them to understand the data. Issues arising from the contextual nature of data and the fundamental challenges of transferring contextual information to data reusers exist across disciplines (e.g., Berg & Goorman, 1999; Cragin & Shankar, 2006; Faniel et al., 2013; Jirotka et al., 2005). Documentation can play an important role in transferring contextual information and supporting data reuse, but reusers reported different experiences working with documentation and of its usefulness (Borgman, 2007; Faniel et al., 2013). Markus (2001) differentiates documentation for oneself, similar others, and dissimilar others and argues that the level of detail and types of contextual information included in the documentation should be different depending on the intended users. According to Niu (2009), documentation for quantitative data in social science tends to be better than that of other kinds of data.

Several studies have demonstrated that human interactions also play an important role in data reuse. Data reusers often search for additional information when documentation is insufficient, consulting various sources, including data producers and experts (Birnholtz & Bietz, 2003; Bishop, 2009; Faniel et al., 2013; Markus, 2001; McCall & Appelbaum, 1991). Yoon (2017) found that social science data reusers also sought external help from data reuser groups, repository staff, and data producers when they encountered problems. Faniel, Kriesberg, and Yakel (2012) reported that human scaffolding, particularly the use of faculty advisors, was an effective technique for novice social science data reusers to manage complex issues that arose during data reuse.

While these studies contribute to the understanding of data reuse practices in the social sciences, explicating relevant factors in data reuse and explaining social science data reuse as compared to other disciplines, only a few studies try to explain or understand data reusers' behaviors from theoretical perspectives or using a theoretical model (e.g., Yoon, 2017). This study develops an integrated theoretical framework to explore social scientists' data reuse behaviors.

4. Conceptual development

4.1. Theory of planned behavior (TPB) and technology acceptance model (TAM)

This study employs an integrated theoretical framework combining the theory of planned behavior (TPB) and the technology acceptance model (TAM). TPB is a social psychology theory linking individuals' behaviors to their attitude toward the behavior, subjective norms, and perceived behavioral controls involved in the behaviors (Ajzen, 1991; Fishbein & Ajzen, 1975). In TPB, attitude, subjective norm, and behavioral control are determined by an individual's attitudinal, normative, and control beliefs (Ajzen, 1991; Fishbein & Ajzen, 1975). Those attitudinal, normative, and control beliefs are formed by a person's fundamental thoughts or views on the results of a certain behavior.

TPB was used in this study to understand social scientists' data reuse behaviors by considering (1) their attitudes toward data reuse, i.e., their overall evaluation of reusing others' data, (2) social scientists' subjective norms of data reuse, i.e., the community expectations among social scientists about data reuse, and (3) perceived behavioral controls (or, resource facilitating conditions) such as data repositories or institutional supports. TPB uses intentions as a proxy for actual behaviors; TPB explains how attitudes, subjective norms, and perceived behavioral controls all influence individuals' intentions to engage in a certain behavior.

Although TPB provides a fundamental theoretical framework for understanding how human behavior is influenced by one's attitude, subjective norm, and behavioral control factors, TPB does not reveal what specific beliefs influence that attitude, norm, or behavior control. Therefore, this study also used TAM to better account for social scientists' perceptions of data reuse. Since the research construct of “attitude toward data reuse” may be influenced by diverse attitudinal beliefs, this research employed TAM to explain specific perceptions toward data reuse. TAM considers utility and effort expectancies, such as perceived usefulness and perceived ease of use, to explain people's intentions to adoption a technology. Integrating TAM with TPB makes sense, as TAM can provide two important research constructs, perceived usefulness of data reuse and perceived effort (c.f., perceived ease of use), involved in data reuse. The TAM construct of “perceived ease of use” was adapted to “perceived effort” to better deliver the idea of effort expectancy involved in data reuse. The utility and effort expectancies captured by TAM are social scientists' perception that data reuse will both benefit them, and be worth their effort. In addition to perceived usefulness and perceived effort, this research integrated the perceived concern involved in data reuse to better explain social scientists' attitude toward data reuse. Perceived concern comprises potential risks involved in data reuse. Social scientists' may not want to reuse other scientists' data if they might infringe upon copyright or not be able to publish their research based on the reused data.

Given the complexity of data reuse intention and behavior, neither theory is sufficient on its own. The integration of TPB and TAM provides a necessary framework by (1) providing a fundamental theoretical model considering attitude, subjective norm, and behavioral control all together (TPB), and (2) suggesting specific attitudinal beliefs which mediate social scientists' attitudes toward data reuse, thereby changing their behavioral intentions.

4.2. Research model and hypotheses development

Based on the integrated theoretical framework above, a research model for social scientists' data reuse was designed. Perceptions were tested, including perceived usefulness, concern, effort involved in data reuse and the influences of these perceptions on social scientists' attitudes toward data reuse. The “subjective norm of data reuse” was also included as part of TPB's subjective norm construct. In TPB, subjective norm is believed to influence both people's behavioral intentions and their attitudes toward a certain behavior (Ajzen, 1991; Ajzen & Fishbein, 2005). The existence of a data repository was also included as an external behavioral control factor which, under TPB, would have a direct influence on people's behavioral intentions and effort expectancy in data reuse.

4.2.1. Perceived usefulness

“Perceived usefulness” refers to the degree to which social scientists believe they would benefit by reusing other researchers' data. In TAM, perceived usefulness is one of the constructs for determining system use; it is understood as the degree to which a person believes using a system will increase relevant job performance (Davis, 1989). In data reuse literature, usefulness is mostly relevant to the benefits that researchers find from reusing existing data, including the increase in research performance and effectiveness. Researchers' decision to reuse data was primarily based on the perceived benefits or the perception that the data met the researchers' needs (Niu, 2009; Yoon, 2015). Pienta, Alter, and Lyle (2010) also considered research productivity as one of the drivers of data reuse. Data reusers are aware that reusing data can increase their research productivity (e.g., increased number of publications, reduced time for data collection). Thus, the perceived usefulness of data reuse should improve social scientists' attitudes toward data reuse.

H1. Perceived usefulness positively affects a social scientist's attitude toward data reuse.

4.2.2. Perceived concern

“Perceived concern” refers to the degree to which social scientists believe data reuse would involve possible risks, such as fewer publication opportunities, misinterpretation of data, or copyright infringement. In the context of social science data reuse, researchers expressed concern about data reuse because they had experienced challenges when trying to publish articles reusing existing data (Yoon, 2014b). As Martin (1995) and Goodwin (2012) suggest, reused data might be perceived as less acceptable or less valuable in some social science disciplines. The possibility of misrepresentation arising from missing information or missing context was also a major concern, not only for data producers but also for data reusers, particularly qualitative researchers (Niu & Hedstrom, 2008; Yoon, 2014b). Perceived concern comprises potential risks involved in data reuse, and TAM was extended by later studies including a perceived risk construct (Featherman & Pavlou, 2003; Wu & Wang, 2005). Given these potential problems, the perceived concern involved in data reuse could lead social scientists to hold negative attitudes about reusing other scientists' data.

H2. Perceived concern negatively affects a social scientist's attitude toward data reuse.

4.2.3. Perceived effort

“Perceived effort” refers to the degree to which social scientists believe that data reuse would require time and energy in order to acquire other scientists' data and to process that data. In TAM, Davis (1989) used “ease of use” as a construct, referring to the degree to which a person believes using a system will be effort-free. “Perceived effort”, rather than “perceived ease of use,” emphasizes the effort expectancy involved in data reuse as compared to the effortlessness or ease of data reuse that a “perceived ease of use” construct would imply (Venkatesh, Morris, Davis, & Davis, 2003). Data reuse literature suggests that when data reusers spend significant time actually using existing data (Rolland & Lee, 2013; Zimmerman, 2008), and less time acquiring and processing data, their level of satisfaction with data reuse improves (Faniel et al., 2016). Thus, it is likely that a greater perceived effort to reuse data would negatively influence social scientists' attitudes toward data reuse and their intention to reuse other scientists' data.

H3. Perceived effort negatively affects a social scientist's attitude toward data reuse.

H4. Perceived effort negatively affects a social scientist's intention to reuse other scientists' data.

4.2.4. Subjective norm of data reuse

The "subjective norm of data reuse" refers to the degree to which social scientists consider data reuse a prevalent research practice in their research communities. Not many data reuse studies have examined the impact of disciplinary norms on data reuse, but previous studies that focused on a specific discipline or across different disciplines suggested that reuse practices were varied and discipline-specific (e.g., Birnholtz & Bietz, 2003; Carlson & Anderson, 2007; Faniel et al., 2013; Rolland & Lee, 2013). This may imply that the norm in a discipline would influence reuse behaviors. In addition, Yoon's (2014b) study indicated that the lack of data reuse norms in certain social science disciplines negatively influences researchers' data reuse behaviors. Thus, the subjective norm of data reuse would positively influence social scientists' attitudes toward data reuse and encourage their intentions to reuse other scientists' data.

H5. Subjective norm of data reuse positively affects a social scientist's attitude toward data reuse.

H6. Subjective norm of data reuse positively affects a social scientist's intention to reuse other scientists' data.

4.2.5. Attitude toward data reuse

"Attitude toward data reuse" refers to the degree to which social scientists believe data reuse is good. The "attitude" in TPB is usually understood as a summary evaluation of a certain behavior, and TPB shows that the attitude toward a certain behavior strongly explains the intention to conduct the behavior (Ajzen, 1991; Ajzen & Fishbein, 2005; Fishbein & Ajzen, 1975). A good number of studies of knowledge sharing have found that positive attitudes toward sharing knowledge lead to positive intentions to share knowledge (Cho, Chen, & Chung, 2010; He & Wei, 2009; Hsu & Chiu, 2004). Thus, a positive attitude toward data reuse would encourage social scientists' intentions to reuse other scientists' data.

H7. Positive attitude toward data reuse positively affects a social scientist's intention to reuse other scientists' data.

4.2.6. Availability of data repository

The availability of a data repository is an important factor influencing social scientists' data reuse. The resource-facilitating conditions found in data repositories can be considered external behavioral controls in TPB (Ajzen, 1991; Ajzen & Fishbein, 2005). Data repositories can reduce perceived effort, enhance positive attitudes, and influence the actual behaviors of social scientists regarding data reuse. In the social science context, data repositories play an important role in data sharing and reuse, although data reuse also still occurs through person-to-person exchange and interaction (e.g., Faniel & Jacobsen, 2010; Yoon, 2014a). Social science data repositories contribute to data reuse, not just through easy access to data, but also by providing value-added services which ensure current and future use of that data, such as managing provenance, correcting errors, and providing supporting documentation, all of which reduce the effort expectancy required for data reuse (Daniels et al., 2012; Fear & Donaldson, 2012). Data repositories also help data reusers trust data because of the repositories' function, reputation, and structure, which encourages researchers to reuse data from repositories (Yakel, Faniel, Kriesberg, & Yoon, 2013; Yoon, 2014a). Thus, the availability of a data repository would reduce social scientists' perceived effort and positively influence their intentions to reuse other scientists' data.

H8. The availability of a data repository negatively affects a social scientist's perceived effort.

H9. The availability of a data repository positively affects a social scientist's intention to reuse other scientists' data.

4.2.7. Intention to reuse data

Data sharing and reuse are not yet well-established research practices in most social science disciplines; because of that limitation, this research measured social scientists' "intentions" to reuse other researchers' data rather than their actual data reuse behaviors. However, while this choice is a practical one, it is nonetheless firmly grounded. In prior knowledge sharing studies, measures of people's intention to engage in a particular behavior serve as a proxy for the actual behavior (Bock, Zmud, Kim, & Lee, 2005; Hsu & Lin, 2008; Kuo & Young, 2008; Lin, 2006). Furthermore, prior studies employing TPB as their theoretical framework support the strong correlation between intentions and actual behaviors (Chen & Chen, 2009; Kuo & Young, 2008; Tsai & Cheng, 2010); therefore, this research is well founded in its use of "intentions" to reuse data as its main outcome variable, for both methodological and theoretical reasons.

Fig. 1 places the hypotheses in the context of the research model incorporating the constructs described above.

5. Research method

Survey methodology was used to evaluate social scientists' data reuse based upon the research constructs in the model. Surveys are

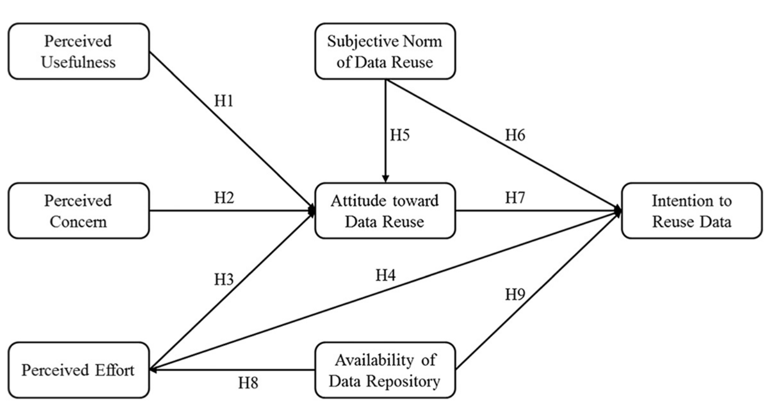

Fig. 1. Research model and hypotheses (H) for social scientists' data reuse intentions.

useful in eliciting people's perceptions of particular behaviors and their intentions to perform those behaviors (Ajzen & Fishbein, 2005; Creswell, 2008).

5.1. Population and sampling

The Community of Science (COS) Scholars database was the major recruitment source. COS includes social science faculty members, postdoctoral researchers, and graduate student researchers in U.S. academic institutions; 96,121 social scientists from the United States were registered in the database (as of September 4th, 2015). 2193 social science scholars were randomly selected and invited to participate. Since structural equation modeling (SEM) was used for data analysis method, a minimum of 200 responses were required to appropriately evaluate the model (Chin, 1998; Fornell & Larcker, 1981).

Compared to prior data reuse studies (Faniel et al., 2016; Yoon, 2014b), this research considered an inclusive range of social science researchers in terms of their data reuse experience (e.g., experienced data reusers, potential data reusers, and even those who never want to reuse data). The survey is also designed to incorporate the wide variety of social science researchers who may have different perceptions (e.g., subjective norm of data reuse) and intentions (e.g. potential data reusers vs those who may not intend to reuse data). Measuring data reuse intentions as the proxy of actual reuse behaviors allowed the researchers to identify important factors which might not have surfaced had the study been limited to subjects with actual data reuse experience.

5.2. Measurement of constructs

The survey questionnaire was designed to measure the research constructs.[1] A total of 21 measurement items for 7 research constructs were adapted from prior studies: perceived usefulness (Davis, 1989), perceived concern (Featherman & Pavlou, 2003), perceived effort (Davis, 1989), attitude toward data reuse (Ajzen & Fishbein, 2005), subjective norm of data reuse (Ajzen & Fishbein, 2005; Scott, 2001), and availability of data repository (Venkatesh, Morris, Davis, & Davis, 2003). In addition, the measurement items for intentions to reuse other scientists' data were adopted from Ajzen and Fishbein's (2005) study. 5-point Likert scales ranging from "strongly agree" to "strongly disagree" were used for the most of survey questions.

5.3. Survey administration

Before distributing the survey, permission was obtained from ProQuest Pivot to use the COS Scholars database. The online survey included a brief description of the study, the study purpose, and the measurement items for each construct (Appendix A). The survey consisted of 27 items, including some demographic questions. On October 5, 2015, the first invitation emails were sent to the 2193 social scientists who were randomly selected from the COS Scholars database. Only one reminder was sent, on November 10, 2015, before the survey was closed on November 30, 2015. Of the 2193 messages sent, 234 messages were returned because of invalid email addresses or spam filters in email servers. A total of 1959 messages were delivered to potential survey participants, and 292 valid responses were received with < 5% of missing values. The response rate was 14.91%. These 292 responses were used for the final data analysis. Only 9 responses out of 292 responses had < 5% of missing values, and mean replacement method was used to treat these missing data in the final data analysis.

Table 1
Demographics.

Demographic category		*n*	%
Gender	Male	172	58.9
	Female	114	39.0
	Missing	6	2.1
Age	25–34	21	7.2
	35–44	61	20.9
	45–54	78	26.7
	55–64	76	26.0
	65 +	55	18.8
	Missing	1	0.3
Ethnic	Caucasian	234	80.1
	Asian/Pacific Islander	18	6.2
	Other/multi-racial	14	4.8
	Hispanic	12	4.1
	Black/African-American	5	1.7
	Native American/Alaska Native	3	1.0
	Missing	6	2.1
Education	PhD/doctoral degree	266	91.1
	Master's degree	20	6.8
	Bachelor's degree	2	0.7
	Missing	4	1.4
Status	Tenured	166	56.8
	Not on tenure track	70	24.0
	Retired	33	11.3
	On tenure track	16	5.5
	Missing	7	2.4
Position	Full professor	98	33.6
	Associate professor	76	26.0
	Professor emeritus	28	9.6
	Assistant professor	23	7.9
	Researcher	13	4.5
	Lecturer/instructor	11	3.8
	Graduate student	11	3.8
	Post-doctoral fellow	6	2.1
	Professor of practice	1	0.3
	Other	24	8.2
	Missing	1	0.3
Total		292	100

5.4. Demographics of participants

The survey participants were mixed in gender (male: 58.9%; female: 39.0%) and age (from 20s to 60s). Caucasians (80.1%) were the most dominant group of respondents. Most respondents were researchers with a Ph.D. degree (91.1%); and more than half of respondents were also tenured or tenure track faculty (62.3%). Their positions were varied but 77.1% held some rank of professorship (from assistant, associate, and full to emeritus). Table 1 summarizes the demographic data.

The survey participants were from diverse social science disciplines (Table 2). Almost a third (30.1%) were in areas of psychology, including clinical psychology (5.1%), non-clinical psychology (5.1%), combined psychology (3.1%), and other areas (16.80%). Sociology (13.0%), anthropology (12.3%), and political science (9.6%) were next most frequently represented.

6. Data analysis and results

Partial least square (PLS) based SEM technique was the primary data analysis method using SmartPLS 2.0 (Ringle, Wende, & Will, 2005). Following the two-stage approach proposed by Anderson and Gerbing (1988), we first tested a measurement model was first tested to examine the reliability and validity of each research construct, then, the structural model was followed to evaluate the research model and hypotheses.

[1] Both survey data and instrument have been made publicly available via Open ICPSR and can be accessed at http://www.openicpsr.org/openicpsr/project/100404/view.

Table 2
Academic disciplines.

Disciplines	*n*	%
Psychology, other	49	16.8
Sociology	38	13.0
Anthropology	36	12.3
Political science	28	9.6
Geography	25	8.6
Clinical psychology	15	5.1
Psychology, except clinical	15	5.1
Public administration	14	4.8
Economics	11	3.8
Psychology, combined	9	3.1
Linguistics	4	1.4
Agricultural economics	2	0.7
History and philosophy of science	2	0.7
Social sciences, other	44	15.1
Total	292	100

Table 3
Reliability and validity values.

Variables	Cronbach's α	CR	AVE
Perceived usefulness	0.76	0.87	0.69
Perceived concern	0.80	0.88	0.71
Perceived effort	0.82	0.88	0.72
Norm of data reuse	0.87	0.92	0.80
Attitude toward data reuse	0.83	0.90	0.75
Availability of data repository	0.91	0.96	0.92
Intention to reuse data	0.97	0.98	0.95

6.1. Measurement model

Prior to the actual data analysis, scale assessment was performed to check the reliability of the measurement scales. Cronbach's α, composite reliability (CR), and average variance extracted (AVE) values are used to ensure each research construct was reliable for further data analysis (Table 3). The scale assessment results showed that Cronbach's α, ranging from 0.76 (perceived usefulness) to 0.97 (intention to reuse data), for all of the measured items had more than the acceptable value of 0.70 (Nunnally & Bernstein, 1994). CR values, ranging from 0.87 (perceived usefulness) to 0.98 (intention to reuse data), were also over the acceptable value of 0.70 (Chin, 1998) and AVE values, ranging from 0.69 (perceived usefulness) and 0.95 (intention to reuse), were all within the acceptable value of 0.50 and above (Fornell & Larcker, 1981).

The measurement models were then examined to ensure the validity of the research constructs, including convergent and discriminant validities. The square roots of the research constructs' AVEs (bolded in Table 4) are greater than the inter-construct correlations (not bolded), indicating reliable convergent and discriminant validity. The measurement model evaluation showed that the measurements for each research construct are reliable and valid for further analysis using structural model evaluation.

The validity of the instrument, including convergent and discriminating validity, was tested by comparing the correlations between one item and the others, and the square root of its AVE. The square roots of AVEs have higher values (convergent validity) than the correlations between the items comprising different constructs (discriminant validity). The square roots of AVEs between items of the same construct ranged from 0.83 (perceived usefulness) to 0.97 (intention to reuse data), which are greater than the inter-item correlation coefficients between items of the different constructs (ranging from − 0.04 to 0.75).

6.2. Structural model

Since the research constructs showed acceptable convergent and discriminant validity, data analysis proceeded and the structural model was estimated using the PLS-SEM technique. The results show that attitudinal, normative, and resource factors all strongly influence social scientists' intentions to reuse others' data. In terms of attitudinal beliefs, perceived usefulness was found to have a significant positive influence on attitude toward data reuse ($\beta = 0.625$, $p < 0.001$), and perceived concern was found to have a significant negative influence on attitude toward data reuse ($\beta = -0.215$, $p < 0.001$). However, perceived effort was not found to have any significant influence on social scientists' attitudes toward data reuse ($\beta = 0.072$, $p > 0.05$). Along with attitudinal belief factors, findings also indicated that social scientists' subjective norms of data reuse had a significant positive influence on their attitudes toward data reuse ($\beta = 0.096$, $p < 0.05$). The perceived usefulness, concern involved in data reuse, and subjective norm of data reuse factors accounted for 60.3% of total variance in attitude toward data reuse ($R^2 = 0.603$).

Social scientists' subjective norms of data reuse showed significant, positive relationships with both attitudes toward data reuse ($\beta = 0.096$, $p < 0.05$) and intention to reuse data ($\beta = 0.169$, $p < 0.01$). The availability of a data repository was found to have a moderate negative influence on the perceived effort involved in data reuse ($\beta = -0.176$, $p < 0.01$); however, it was not found to have any significant direct influence on intention to reuse data ($\beta = 0.091$, $p > 0.05$). Perceived effort, as a behavioral control variable, was found to have a significant negative influence on intention to reuse data ($\beta = -0.153$, $p < 0.001$) directly. Lastly, attitude toward data reuse was found to have a significant positive influence on intention to reuse data ($\beta = 0.467$, $p < 0.001$). The subjective norm of data reuse, attitude toward data reuse, and perceived effort explain 38.4% of total variance in social scientists' intentions to reuse data ($R^2 = 0.384$). Fig. 2 shows the results of hypothesis testing based on the model, and Table 5 summarizes hypothesis testing results.

Table 4
Correlation matrix, square roots of AVEs.

	Perceived usefulness	Perceived concern	Perceived effort	Norm of data reuse	Attitude toward data reuse	Availability of data repository	Intention to reuse data
Perceived usefulness	**0.83**						
Perceived concern	− 0.41	**0.84**					
Perceived effort	− 0.07	0.28	**0.85**				
Norm of data reuse	0.38	− 0.29	− 0.13	**0.89**			
Attitude toward data reuse	0.75	− 0.48	− 0.04	0.39	**0.87**		
Availability of data repository	0.18	− 0.22	− 0.18	0.31	0.19	**0.96**	
Intention to reuse data	0.62	− 0.42	− 0.21	0.40	0.26	0.56	**0.97**

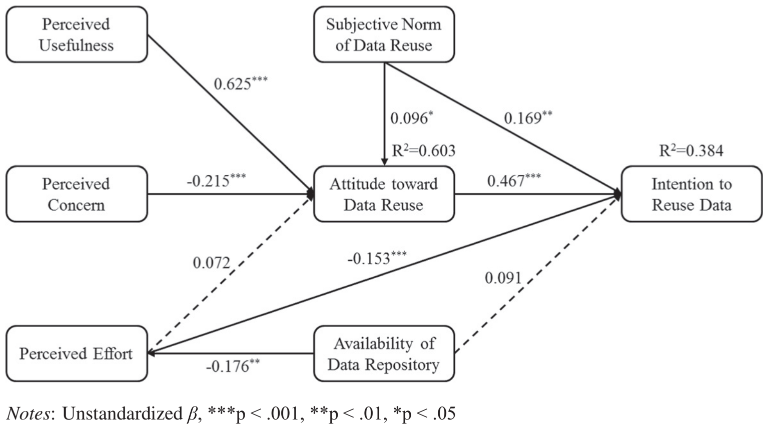

Fig. 2. Hypothesis testing results based on social scientists' data reuse model. Notes: Unstandardized β, $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$.

Table 5
Summary of hypothesis testing results.

Hs	Statements	Result	Beta (p)
H1	Perceived usefulness would positively affects a social scientist's attitude toward data reuse	Supported	0.625***
H2	Perceived concern would negatively affects a social scientist's attitude toward data reuse	Supported	− 0.215***
H3	Perceived effort would negatively affects a social scientist's attitude toward data reuse	NOT Supported	0.072
H4	Perceived effort would negatively affects a social scientist's intention to reuse other scientists' data	Supported	− 0.153***
H5	Subjective norm would positively affects a social scientist's attitude toward data reuse	Supported	0.096*
H6	Subjective norm would positively affects a social scientist's intention to reuse other scientists' data	Supported	0.169**
H7	The availability of data repository would negatively affects a social scientist's perceived effort	Supported	− 0.176**
H8	The availability of data repository would positively affects a social scientist's intention to reuse other scientists' data	NOT Supported	0.091
H9	Attitude toward data reuse would positively affects a social scientist's intention to reuse other scientists' data	Supported	0.467***

*** $p < 0.001$.
** $p < 0.01$.
* $p < 0.05$.

7. Discussion

The integrated theoretical framework explains which attitudinal beliefs (i.e., perceived usefulness and concern) influence social scientists' attitudes toward data reuse based on TAM, and then how those attitudes toward data reuse, along with subjective norms of data reuse and the availability of a data repository, affect social scientists' data reuse intentions, in accordance with TPB.

Attitudes toward data reuse were influenced by perceived usefulness and perceived concern. Social scientists consider reusing others' data because they perceive that doing so would increase their research performance and productivity. However, they are hesitant to reuse others' data when they think doing so could potentially cause problems, such as misrepresentation of data, copyright infringement, and/or fewer publication opportunities. In addition, social scientists develop positive attitudes toward data reuse if they believe that their communities or disciplines have strong norms of data reuse. These findings demonstrate the significance of informing and educating social scientists about the potential and benefits of data while mitigating possible concerns. Since having a positive attitude toward data reuse influences intention to reuse data, educating and informing them would be an important first step toward actual data reuse.

It is worth noting that the established community/disciplinary norms of data reuse were found to have significant positive impacts on both attitudes toward data reuse and intentions to reuse data in this study. When social scientists perceive an expectation of data reuse and acknowledgement by their communities and disciplines that data reuse is common and acceptable, they are more likely to have strong norms of data reuse and to have positive attitudes toward data reuse. This suggests that social science communities need a better and stronger subjective norm of data reuse to lead to more active data reuse behavior among researchers. Not many research studies have reported on what the norm for data reuse in each field of social science is yet, but it is true that some fields have a more active culture of data reuse, where others do not. Even within a particular field, the subjective norm of data reuse may vary depending on the types of research done. These differing norms might well be reflected in differing levels of actual data reuse. Further exploring the data reuse behaviors in a variety of disciplines—both with and without strong norms of data reuse—will help us to better understand the impact of norms in data reuse behaviors and lead to a better understanding of how a community can create a data reuse culture.

Perceived effort of data reuse has a strong negative influence on social scientists' data reuse intentions. In addition, the availability of a data repository was an important resource factor found to have a significant negative relationship with the perceived effort involved in data reuse. Although the availability of a data repository was not found to have a significant relationship with data reuse intentions, it influences data reuse intention indirectly through its impact upon perceived effort. This finding suggests that the availability of a data repository can reduce the effort expectancy involved in data reuse, and lower effort expectancy can increase social scientists' intentions to reuse others' data. When social scientists believe reusing others' data requires too much effort, they are less likely to reuse data. It is important to reduce their effort expectancy by providing resources and support, such as data repositories, to encourage data reuse.

7.1. Theoretical implications

This research has significant theoretical and practical implications. While previous studies on data reuse have empirically investigated researchers' experiences, perceptions, and attitudes regarding data reuse, few have employed a theoretical approach to explain behaviors and to investigate the relationships among various factors that influence data reuse. The present, novel, study offers a theoretical framework integrating TPB and TAM to explain researchers' data reuse behavior. The combination of TPB and TAM provides a more comprehensive theoretical framework for studying social scientists' data reuse, addressing both the social-psychological and practical considerations driving data reuse adoption or rejection by individuals. The results demonstrate that the framework is well supported by the survey data. The combined model using TPB and TAM offers a broader understanding that neither theory could provide independently. TPB alone cannot account for social scientists' data reuse behaviors, and TAM is useful in explaining how perceived usefulness and effort expectancies influence social scientists' attitude toward data reuse and their intentions to reuse others' data. Along with perceived usefulness and effort, perceived concern about data reuse was added to the model, and the relationship between perceived concern and attitude toward data reuse received empirical support from the survey data. This study showed a significant amount of variance in social scientists' attitudes toward data reuse, which is explained by the perceived usefulness of reused data, the concerns involved in data reuse, and communities' subjective norms regarding data reuse. Those diverse factors can improve the explanatory power of research into social scientists' attitude toward data reuse. Furthermore, this study validated the research model based on the survey data with social scientists; the validation of this model points to the potential of its generalizability as a framework for understanding social science data reuse.

7.2. Practical implications

The findings also have several practical implications for stakeholders interested in promoting and facilitating data reuse. First, it is critical to enhance social scientists' awareness of the potential and benefits of data reuse while decreasing their concerns regarding data reuse. While decisions regarding data reuse are ultimately influenced by individual researchers' interests and needs, reaching out to the broader communities of researchers, particularly those who have not reused existing data yet, would help researchers understand how data reuse can improve their research and productivity. While research libraries have started to design and provide instructions for data reuse, they need to be more proactive in informing and educating researchers and addressing researchers' concerns regarding data reuse. Research communities should also be involved in creating workshops or training sessions on data reuse and sharing data reuse experiences with other researchers; in addition to educating researchers on the usefulness of data reuse, education by members of one's own research community will enforce a subjective community norm of data reuse.

Given the centrality of that subjective norm of data reuse, data reuse is not just driven by personal motivation and interests. Rather, data reuse is also driven by the social norms in disciplines or research communities; those norms are best understood as an informal understanding of the codes of conduct and behaviors that individuals can enact (Lapinski & Rimal, 2005). Nurturing a culture of data reuse would take time, especially if a discipline or community has a tradition of not allowing data reuse. Considering that more and more disciplines are becoming open to data reuse (with rationales as to why reusing existing data is methodologically acceptable) and laws, policies, and regulations have been developed which support it, social scientists may need to revisit the research traditions in their disciplines, including addressing methodological and/or ethical concerns, in order to promote a culture of data reuse. Research support from libraries and repositories will also be essential in creating a subjective norm of data reuse in social science disciplines, as each discipline is still subject to the broader norms of its institutions.

The role of data repositories stood out as encouraging data reuse by reducing researchers' efforts in data reuse. As mentioned above, social science data repositories have a long history of serving their designated communities through their value-adding activities, such as access, documentation, data cleaning, version controls, error management, and preservation (Daniels et al., 2012; Yoon, 2014a), which reduce social scientists' effort expectancies regarding data reuse. This study reinforces the importance of data repositories in the data reuse landscape and calls for their more active involvement in data reuse. Despite the work of data repositories in data sharing and reuse, many researchers still reported some common difficulties in reusing data, including: not all data being deposited in repositories, issues with documentation, and even not knowing how to search data (Faniel & Jacobsen, 2010; Sands, Borgman, Wynholds, & Traweek, 2012). While some repositories, like ICPSR, are well known across different social science fields, other repositories may need to reach out to research communities to broadcast their services and datasets. Given that efforts have also been initiated to improve repositories' processes and services for curating data by developing guidelines, best practices, and policies (e.g., DataCite, 2017), repositories will likely continue to facilitate social scientists' data sharing and reuse behaviors.

7.3. Limitations

This study has some methodological limitations. The survey method is limited because it does not capture much contextual information about social scientists' data reuse behaviors. The practices of data reuse can be different in each individual researcher and how researchers interact with what types and formats of data can vary person to person. Follow-up qualitative research would be useful, as it would complement this study, and would offer a more refined understanding of social scientists' data reuse behaviors.

For pragmatic reasons explained above, data reuse intentions were used as a proxy for actual data reuse behaviors. Although many research studies support the use of intentions as proxies for actual behaviors (Chen & Chen, 2009; Kuo & Young, 2008; Tsai & Cheng, 2010), there is a possibility that data reuse intentions may not explain social scientists' actual data reuse behaviors. Because data reuse is a complex and ongoing process in which unexpected challenges may arise (Zimmerman, 2008), expert and novice data reusers' perceptions may vary. Thus, future studies should measure actual behaviors of data reuse rather than intentions to reuse data, to build on the groundwork laid here and provide a richer understanding of data reuse in the social sciences.

8. Conclusion

This study investigated factors that influence social scientists' data reuse behaviors using a theoretical model combining TPB and TAM. The findings of this study make an important contribution to the data reuse literature by providing a broad explanation of the relationships among factors that influence data reuse behaviors. This study suggests that subjective norm of data reuse, attitudinal beliefs (e.g., perceived usefulness and concern) and attitude toward data reuse, perceived effort involved in data reuse, and availability of data repository should be taken into account in any effort to facilitate social scientists' data reuse. Social scientists' data reuse behaviors can be addressed by lessening the concerns and efforts involved in data reuse, as well as emphasizing the usefulness of data reuse, increasing positive norms of data reuse, and providing reachable data repositories. From the standpoint of librarianship, academic libraries can support social scientists' data reuse behaviors by providing data search and management services, consulting on copyright and ethical issues in data reuse, and educating researchers about data reuse practices in the social sciences. The findings of this study can encourage scientific data sharing and reuse in social science communities. This research is a step toward enabling data-intensive research in social sciences by suggesting means of supporting and promoting the reuse of shared and accumulated data sets.

Acknowledgements

The authors would like to acknowledge the ProQuest Pivot for allowing the use of its Community of Scientists Scholar Database in recruiting the survey participants, and thank Darra Hofman for reviewing the paper.

Appendix

Appendix A
Measurement items for research constructs.

Construct	Items	Sources
Perceived usefulness	• Reusing other researchers' data improves the quality of my research • Reusing other researchers' data enhances the effectiveness of my research • Reusing other researchers' data reduces the time/cost/effort I spend on my research	(Bock, Zmud, Kim, & Lee, 2005; Kim & Stanton, 2016; Wasko & Faraj, 2000)
Perceived concern	• If I reuse other researchers' data, I worry that I might misinterpret the data • If I reuse other researchers' data, I worry that I might cause infringement • If I reuse other researchers' data, I worry that I might not publish with that data	(Featherman & Pavlou, 2003; Kim & Stanton, 2016; Pavlou, 2003)
Perceived effort	• Reusing other researchers' data requires time and effort to locate data sets • Reusing other researchers' data requires time and effort to access (or get permission to use) data sets • Reusing other researchers' data requires time and effort to process data sets for a new study	(Davis, 1989; Kim & Stanton, 2016; Thompson, Higgins, & Howell, 1991)
Attitude toward data reuse	• Reusing other researchers' data is valuable • Reusing other researchers' data is desirable • Reusing other researchers' data is pleasant	(Ajzen & Fishbein, 2005)
Subjective norm of data reuse	• In my discipline, it is expected that researchers could reuse other researchers' data • In my discipline, many of researchers currently reuse data • In my discipline, reusing other researchers' data is a common practice	(Kim & Stanton, 2016; Kostova & Roth, 2002; Son & Benbasat, 2007)
Availability of data repository	• In my discipline, data repositories are available for researchers to share data • In my discipline, researchers can easily access data repositories to reuse data	(Kim & Stanton, 2016; Taylor & Todd, 1995; Venkatesh, Morris, Davis, & Davis, 2003)
Intention to reuse data	• I am likely to reuse other researchers' data for my future research • I intend to reuse other researchers' data for my future research • I will try to reuse other researchers' data for my future research	(Ajzen & Fishbein, 2005)

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Process*, *52*(2), 179–211.

Ajzen, I., & Fishbein, M. (2005). The influence of attitudes on behavior. In D. Albarracin, B. T. Johnson, & M. P. Zanna (Eds.), *Handbook of attitudes and attitude change* (pp. 173–221). Mahwah, NJ: Lawrence Erlbaum Associates.

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, *103*(3), 411–423.

Berg, M., & Goorman, E. (1999). The contextual nature of medical information. *International Journal of Medical Informatics*, *56*(1–3), 51–60. http://dx.doi.org/10.1016/s1386-5056(99)00041-6.

Bergman, M. M., & Eberle, T. S. (Eds.). (2005). Qualitative inquiry: Research, archiving, and reuse [Special issue], *Forum: Qualitative Social Research*, *6*(2).

Birnholtz, J. P., & Bietz, M. J. (2003). Data at work: Supporting sharing in science and engineering. In K. Schmidt & M. Pendergast (Chairs), Proceedings of the International ACM Conference on Supporting Group Work (pp. 339-348). New York, NY: ACM. doi: http://dx.doi.org/10.1145/958160.958215.

Bishop, L. (2009). Ethical sharing and reuse of qualitative data. *Australian Journal of Social Issues*, *44*(3), 255–272.

Bock, G. W., Zmud, R. W., Kim, Y. G., & Lee, J. N. (2005). Behavioral intention formation in knowledge sharing: Examining the roles of extrinsic motivators, social-psychological forces, and organizational climate. *MIS Quarterly*, *29*(1), 87–111.

Borgman, C. L. (2007). *Scholarship in the digital age: Information, infrastructure, and the Internet*. Cambridge, MA: MIT Press.

Borgman, C. L. (2012). The conundrum of sharing research data. *Journal of the American Society for Information Science and Technology*, *63*, 1059–1078. http://dx.doi.org/10.1002/asi.22634.

Carlson, S., & Anderson, B. (2007). What are data? The many kinds of data and their implications for data re-use. *Journal of Computer-Mediated Communication*, *12*(2), 635–651.

Chen, I. Y., & Chen, N.-S. (2009). Examining the factors influencing participants' knowledge sharing behavior in virtual learning communities. *Journal of Educational Technology & Society*, *12*(1), 134.

Chin, W. W. (1998). The partial least squares approach to structural equation modeling. *Modern Methods for Business Research*, *295*(2), 295–336.

Cho, H. C., Chen, M. H., & Chung, S. Y. (2010). Testing an integrative theoretical model of knowledge-sharing behavior in the context of Wikipedia. *Journal of the American Society for Information Science and Technology*, *61*, 1198–1212. http://dx.doi.org/10.1002/asi.21316.

Clubb, J. M., Austin, E. W., Geda, C. L., & Traugott, M. W. (1985). Sharing research data in the social sciences. In S. E. Fienberg, M. E. Martin, & M. L. Straf (Eds.), *Sharing research data* (pp. 39–88). Washington, DC: National Academy Press.

Cragin, M. H., & Shankar, K. (2006). Scientific data collections and distributed collective practice. *Computer Supported Cooperative Work*, *15*(2–3), 185–204.

Creswell, J. W. (2008). *Research design: Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, CA: SAGE.

Curty, R. G. (2015). *Beyond "data thrifting": An investigation of factors influencing research data reuse in the social sciences* (Unpublished doctoral dissertation)Syracuse, NY: Syracuse University.

Daniels, M., Faniel, I. M., Fear, K., & Yakel, E. (2012). Managing fixity and fluidity in data

repositories. In J.-E. Mai (Chair), Proceedings of the iConference 2012, February 7–10, 2012, Toronto, Ontario, Canada (pp. 279-286). New York, NY: ACM. doi: http://dx.doi.org/10.1145/2132176.2132212.
DataCite (2017). Our mission. Retrieved from https://datacite.org/mission.html.
Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance in information technology. *MIS Quarterly*, *13*(3), 319–340.
Faniel, I. M., & Jacobsen, T. E. (2010). Reusing scientific data: How earthquake engineering researchers assess the reusability of colleagues' data. *Computer Supported Cooperative Work*, *19*(3–4), 355–375. http://dx.doi.org/10.1007/s10606-010-9117-8.
Faniel, I. M., Kansa, E., Kansa, S. W., Barrera-Gomez, J., & Yakel, E. (2013). The challenges of digging data: study of context in archaeological data reuse. *Proceedings of the ACM/IEEE-CS Joint Conference on Digital Libraries*. *13*. *Proceedings of the ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 295–304). . http://dx.doi.org/10.1145/2467696.2467712.
Faniel, I. M., Kriesberg, A., & Yakel, E. (2012). Data reuse and sensemaking among novice social scientists. *Proceedings of the American Society for Information Science and Technology Annual Meeting*, *49*(1), 1–10.
Faniel, I. M., Kriesberg, A., & Yakel, E. (2016). Social scientists' satisfaction with data reuse. *Journal of the Association for Information Science and Technology*, *67*, 1404–1416. http://dx.doi.org/10.1002/asi.23480.
Fear, K. (2013). *Measuring and anticipating the impact of data reuse* (Unpublished doctoral dissertation)Ann Arbor, MI: University of Michigan.
Fear, K., & Donaldson, D. R. (2012). Provenance and credibility in scientific data repositories. *Archival Science*, *12*, 319–339.
Featherman, M. S., & Pavlou, P. A. (2003). Predicting e-services adoption: A perceived risk facets perspective. *International Journal of Human-Computer Studies*, *59*, 451–474. http://dx.doi.org/10.1016/s1071-5819(03)00111-3.
Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior*. Reading, MA: Addison-Wesley.
Fornell, C., & Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: Algebra and statistics. *Journal of Marketing Research*, *18*(3), 382–388.
Goodwin, J. (2012). *The secondary analysis of qualitative data*. Thousand Oaks, CA: Sage Publishing.
He, W., & Wei, K. K. (2009). What drives continued knowledge sharing? An investigation of knowledge-contribution and -seeking beliefs. *Decision Support Systems*, *46*, 826–838. http://dx.doi.org/10.1016/j.dss.2008.11.007.
Hedrick, T. E. (1988). Justifications for the sharing of social science data. *Law and Human Behavior*, *12*(2), 163–171. http://dx.doi.org/10.1007/bf01073124.
Hsu, C.-L., & Lin, J. C.-C. (2008). Acceptance of blog usage: The roles of technology acceptance, social influence and knowledge sharing motivation. *Information & Management*, *45*(1), 65–74. http://dx.doi.org/10.1016/j.im.2007.11.001.
Hsu, M. H., & Chiu, C. M. (2004). Predicting electronic service continuance with a decomposed theory of planned behaviour. *Behaviour & Information Technology*, *23*, 359–373. http://dx.doi.org/10.1080/01449290410001669969.
Inter-university Consortium for Political and Social Research (2016). *Glossary of social science terms*. Retrieved from https://www.icpsr.umich.edu/icpsrweb/ICPSR/support/glossary.
Jirotka, M., Procter, R., Hartswood, M., Slack, R., Simpson, A., Coopmans, C., ... Voss, A. (2005). Collaboration and trust in healthcare innovation: The eDiaMoND case study. *Computer Supported Cooperative Work*, *14*, 369–398.
Kim, Y., & Stanton, J. M. (2016). Institutional and individual factors affecting scientists' data-sharing behaviors: A multilevel analysis. *Journal of the Association for Information Science and Technology*, *67*, 776–799. http://dx.doi.org/10.1002/asi.23424.
Kostova, T., & Roth, K. (2002). Adoption of an organizational practice by subsidiaries of multinational corporations: Institutional and relational effects. *Academy of Management Journal*, *45*(1), 215–233.
Kuo, F. Y., & Young, M. L. (2008). Predicting knowledge sharing practices through intention: A test of competing models. *Computers in Human Behavior*, *24*, 2697–2722. http://dx.doi.org/10.1016/j.chb.2008.03.015.
Lapinski, M. K., & Rimal, R. N. (2005). An explication of social norms. *Communication Theory*, *15*(2), 127–147. http://dx.doi.org/10.1093/ct/15.2.127.
Lin, H.-F. (2006). Impact of organizational support on organizational intention to facilitate knowledge sharing. *Knowledge Management Research & Practice*, *4*(1), 26–35.
Markus, M. L. (2001). Toward a theory of knowledge reuse: Types of knowledge reuse situations and factors in reuse success. *Journal of Management Information Systems*, *18*(1), 57–93.
Martin, P. A. (1995). More replication studies needed. *Applied Nursing Research*, *8*(2), 102–103.
McCall, R. B., & Appelbaum, M. I. (1991). Some issues of conducting secondary analyses. *Developmental Psychology*, *27*(6), 911–917. http://dx.doi.org/10.1037/0012-1649.27.6.911.
Niu, J. (2009). Overcoming inadequate documentation. *Proceedings of the American Society for Information Science and Technology Annual Meeting* (pp. 46–). . Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/meet.2009.145046024/pdf.
Niu, J., & Hedstrom, M. (2008, October). Documentation evaluation model for social science data. *Paper presented at the American Society for Information Science and Technology Annual Meeting, October 24–29, olumbus, Ohio*. Retrieved from https://deepblue.lib.umich.edu/handle/2027.42/63090.
Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.
Pavlou, P. A. (2003). Consumer acceptance of electronic commerce: Integrating trust and risk with the technology acceptance model. *International Journal of Electronic Commerce*, *7*(3), 101–134.
Pienta, A. M., Alter, G. C., & Lyle, J. A. (2010, April). The enduring value of social science research: The use and reuse of primary research data. *Paper presented at the Organisation, Economics and Policy of Scientific Research Workshop, Torino, Italy, April 2010*. Retrieved from http://deepblue.lib.umich.edu/handle/2027.42/78307.
The book of the Bremen Workshop [Special issue]. In K. B. Rasmussen (Vol. Ed.), *IASSIST Quarterly*. *34 (3 & 4)*.
Ringle, C. M., Wende, S., & Will, S. (2005). *SmartPLS 2.0 (M3) Beta*. Retrieved from http://www.smartpls.de.
Rolland, B., & Lee, C. P. (2013). Beyond trust and reliability: Reusing data in collaborative cancer epidemiology research. *ACM Conference on Computer Supported Cooperative Work and Social Computing*, *16*, 435–444. http://dx.doi.org/10.1145/2441776.2441826.
Sands, A., Borgman, C. L., Wynholds, L., & Traweek, S. (2012). Follow the data: How astronomers use and reuse data. *Proceedings of the American Society for Information Science and Technology*, *49*(1), 1–3.
Scott, W. R. (2001). *Institutions and organizations* (2nd ed.). Thousand Oaks, CA: Sage Publications.
Son, J. Y., & Benbasat, I. (2007). Organizational buyers' adoption and use of B2B electronic marketplaces: Efficiency- and legitimacy-oriented perspectives. *Journal of Management Information Systems*, *24*(1), 55–99. http://dx.doi.org/10.2753/mis0742-1222240102.
Taylor, S., & Todd, P. A. (1995). Understanding information technology usage: A test of competing models. *Information Systems Research*, *6*(2), 144–176.
Thompson, R. L., Higgins, C. A., & Howell, J. M. (1991). Personal computing: Toward a conceptual model of utilization. *MIS Quarterly*, *15*(1), 125–142.
Tsai, M. T., & Cheng, N. C. (2010). Programmer perceptions of knowledge-sharing behavior under social cognitive theory. *Expert Systems with Applications*, *37*, 8479–8485. http://dx.doi.org/10.1016/j.eswa.2010.05.029.
Van House, N. A. (2002). Digital libraries and practices of trust: Networked biodiversity information. *Social Epistemology*, *16*(1), 99–114.
Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, *27*(3), 425–478.
Wasko, M. M., & Faraj, S. (2000). “It is what one does”: Why people participate and help others in electronic communities of practice. *Journal of Strategic Information Systems*, *9*(2–3), 155–173.
Wu, J. H., & Wang, S. C. (2005). What drives mobile commerce? An empirical evaluation of the revised technology acceptance model. *Information & Management*, *42*(5), 719–729. http://dx.doi.org/10.1016/j.im.2004.07.001.
Yakel, E., Faniel, I. M., Kriesberg, A., & Yoon, A. (2013). Trust in digital repositories. *International Journal of Digital Curation*, *8*(1), 143–156.
Yoon, A. (2014a). End users' trust in data repositories: Definition and influences on trust development. *Archival Science*, *14*(1), 17–34.
Yoon, A. (2014b). “Making a square fit into a circle”: Researchers’ experiences reusing qualitative data. *Proceedings of the American Society for Information Science and Technology Annual Meeting*, *51*. http://dx.doi.org/10.1002/meet.2014.14505101140.
Yoon, A. (2015). *Data reuse and users' trust judgments: Toward trusted data curation* (Unpublished doctoral dissertation)Chapel Hill, NC: University of North Carolina at Chapel Hill.
Yoon, A. (2016, March). Visible evidence of invisible quality dimensions and the role of data management. *Paper presented at the iConference, Philadelphia, PA, March 20–23, 2016*.
Yoon, A. (2017). Data reusers' trust development. *Journal of the Association for Information Science and Technology*, *68*, 946–956. http://dx.doi.org/10.1002/asi.23730.
Zimmerman, A. (2008). New knowledge from old data: The role of standards in the sharing and reuse of ecological data. *Science Technology & Human Values*, *33*, 631–652. http://dx.doi.org/10.1177/0162243907306704.

Ayoung Yoon is an assistant professor at the School of Informatics and Computing, Department of Library and Information Science, Indiana University Indianapolis. She earned her PhD from the University of North Carolina at Chapel Hill. Her research areas include data curation, data sharing and reuse in the context of open data and citizen science. Her articles have been published in *Archival Science*, *International Journal of Digital Curation*, and *Journal of the Association for Information Science and Technology*. She is also an RDA/US data share fellow for 2016–2017.

Youngseek Kim is an assistant professor in the School of Information Science at the University of Kentucky. He completed his PhD at Syracuse University, and he received the Eugene Garfield Doctoral Dissertation Award from the Association for Library and Information Science Education and the Emerald/EFMD Outstanding Doctoral Research Award (Highly Commended) in Information Science. His current research efforts in e-science explore the nature of scientists' data practices, focusing on data sharing and reuse, the education of e-science professionals, and adoption and use of cyberinfrastructure. His articles appear in *International Journal of Information Management, Journal of the Association for Information Science and Technology*, *Journal of Information Science*, and *Library and Information Science Journal*.